

Release 2.1D John F. Collins, Biocomputing Research  
Unit.  
Copyright (c) 1993, 1994, 1995 University of  
Edinburgh, U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using  
Smith-Waterman algorithm

Run on: Wed Nov 1 08:32:29 1995; MasPar time 10.28  
Seconds

476.357 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-54  
Description: (1:771) from US08121713B.pep  
Perfect Score: 5765  
Sequence: 1  
MGWLTRIVCLFWGVLLTARA.....ENKKGRNRRTHEFERAPRSV 771

Scoring table: PAM 150  
Gap 11

Searched: 53402 seqs, 6354270 residues

Database: a-geneseq  
1 a-gen1  
2 a-gen2  
3 a-gen3  
4 a-gen4  
5 a-gen5  
6 a-gen6  
7 a-gen7  
8 a-gen8  
9 a-gen9  
10 a-gen10

Statistics: Mean 37.851; Variance 162.907; scale 0.232

Pred. No. is the number of results predicted by chance to  
have a  
score greater than or equal to the score of the result  
being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query					
No.	Score	Match	Length	DB	ID	Description
Pred. No.						

-----

-----						
1	108	1.9	485	10	R55585	AmEPV NPH-1.
2.98e+00						
2	102	1.8	240	3	R15263	Extracellular domain
7.88e+00						
3	101	1.8	355	3	R15264	Extracellular domain
9.25e+00						
4	101	1.8	526	3	R15266	Clone pTB1228-encoded
9.25e+00						
5	101	1.8	643	3	R15267	Clone pTB1229-encoded
9.25e+00						
6	101	1.8	652	3	R15269	Clone pTB1283-encoded
9.25e+00						
7	101	1.8	821	4	R21080	flg receptor protein.
9.25e+00						
8	101	1.8	679	8	R41517	K-sam.
9.25e+00						
9	100	1.7	769	3	R15268	Clone pTB1284-encoded
1.08e+01						
10	98	1.7	726	2	R10933	KGF receptor.
1.49e+01						
11	98	1.7	1365	6	R36780	KRE5.
1.49e+01						
12	98	1.7	726	3	R14280	Murine KGF receptor.
1.49e+01						
13	92	1.6	106	5	R26953	Human T lymphocyte re
3.77e+01						
14	92	1.6	1257	9	R46627	Neurocan core protein
3.77e+01						
15	91	1.6	468	1	P90525	B cell stimulating fa
4.39e+01						
16	91	1.6	460	4	R22616	IL-6R for soluble IL-
4.39e+01						
17	90	1.6	691	7	R38735	Beta-galactosidase.
5.11e+01						
18	90	1.6	403	1	P91165	38 kd regression-asso
5.11e+01						
19	90	1.6	403	8	R40855	38kd regression assoc
5.11e+01						
20	89	1.5	498	3	R13270	Lymphocyte Activation
5.94e+01						
21	89	1.5	380	3	R13272	Polyhedrin-soluble LA
5.94e+01						
22	88	1.5	131	5	R25579	Insulin precursor MI3
6.91e+01						
23	87	1.5	118	10	R57330	NMDAR1-3C residues 80
8.02e+01						
24	87	1.5	102	10	R57333	NMDAR1-4 residues 803
8.02e+01						
25	87	1.5	922	10	R55531	Human NMDA R1d recept
8.02e+01						
26	87	1.5	118	10	R57329	NMDAR1-3A residues 80
8.02e+01						
27	87	1.5	816	3	R14444	A.faecalis penicillin

8.02e+01							
28	87	1.5	911	2	R10333	Deduced sequence of t	
8.02e+01							
29	87	1.5	938	10	R55532	Human NMDA R1a recept	
8.02e+01							
30	87	1.5	920	9	R49043	NMDA receptor channel	
8.02e+01							
31	87	1.5	920	9	R49044	NMDA receptor channel	
8.02e+01							
32	86	1.5	386	9	R51061	Sequence encoded by t	
9.30e+01							
33	86	1.5	283	8	R45455	Theileria sergenti ma	
9.30e+01							
34	86	1.5	886	1	P80345	Sequence encoded by S	
9.30e+01							
35	85	1.5	344	1	P90528	B cell stimulating fa	
1.08e+02							
36	85	1.5	460	2	P81061	Sequence of rhinoviru	
1.08e+02							
37	85	1.5	2164	1	P80131	Peptides translated f	
1.08e+02							
38	85	1.5	3685	1	P90290	Human Duchenne muscul	
1.08e+02							
39	85	1.5	468	3	P90284	Sequence of a recepto	
1.08e+02							
40	85	1.5	323	1	P90527	B cell stimulating fa	
1.08e+02							
41	85	1.5	3685	3	P90373	Sequence encoded by h	
1.08e+02							
42	85	1.5	2164	2	P81045	Sequence of the viral	
1.08e+02							
43	85	1.5	110	5	R26951	Human T lymphocyte re	
1.08e+02							
44	85	1.5	247	8	R42382	H. somnus lppA.	
1.08e+02							
45	85	1.5	265	3	R12844	HTLV-1 protein expres	
1.08e+02							

#### ALIGNMENTS

RESULT 1

ID R55585 standard; Protein; 485 AA.

AC R55585;

DT 18-JAN-1995 (first entry)

DE AmEPV NPH-1.

KW Spheroidin; gene expression; vector; insect cell culture;

KW mammal cell culture; AmEPV; NPH-1;

KW nucleoside-triphosphate-phosphohydrolase.

OS Amsacta moorei entomopoxvirus.

PN WO9413812-A.

PD 23-JUN-1994.

PF 07-DEC-1993; U11907.

PR 07-DEC-1992; US-991867.

PA (UYFL ) UNIV FLORIDA.  
 PI Gruidl ME, Hall RL, Moyer RW;  
 DR WPI; 94-217887/26.  
 PT New entomopoxvirus polynucleotide sequences, proteins and  
 vectors  
 PT - are used for expression of heterologous proteins in both  
 insect  
 PT and mammalian host cells  
 PS Disclosure; Page 85-86; 118pp; English.  
 CC The AmEPV spheroidin gene was isolated and sequenced.  
 Mapping  
 CC showed the gene to be located at the 3' end of a nucleoside-  
 CC triphosphate-phosphohydrolase gene (NPH-1). The sequence of  
 the  
 CC NPH-1 gene is given in Q66824, and its encoded amino acid  
 sequence  
 CC in R55585. The spheroidin gene can be used as the location  
 for  
 CC insertion of heterologous DNA in insect and mammalian  
 expression  
 CC systems.  
 SQ Sequence 485 AA;

DB 10; Score 108; Match 23.9%; QryMatch 1.9%; Pred. No.  
 2.98e+00;  
 Matches 16; Conservative 26; Mismatches 21; Indels  
 4; Gaps 4;

Db 241  
 sykrlieadsltetnyidgyakknifyhniimsdeqsklynmaekydyktelggglktmrr 300  
 ||| ::|::: | : :| |: : || ::: :|:|:| | : :|  
 :| ::  
 Qy 35  
 SYKEMLESNNVI-T-F-NGLANSSS-YHTFLLDEERSRLYVGAKDHIFSFDLVNIKDFQK 90

Db 301 lissfaf 307  
 :: : ::  
 Qy 91 IVWPVSY 97

Search completed: Wed Nov 1 08:32:56 1995  
 Job time : 27 secs.

Release 2.1D John F. Collins, Biocomputing Research  
Unit.  
Copyright (c) 1993, 1994, 1995 University of  
Edinburgh, U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using  
Smith-Waterman algorithm

Run on: Wed Nov 1 08:31:01 1995; MasPar time 27.72  
Seconds

663.261 Million cell

updates/sec  
Tabular output not generated.

Title: >US-08-121-713B-54  
Description: (1:771) from US08121713B.pep  
Perfect Score: 5765  
Sequence: 1  
MGWLTRIVCLFWGVLLTARA.....ENKKGRNRRTHEFERAPRSV 771

Scoring table: PAM 150  
Gap 11

Searched: 78488 seqs, 23849247 residues

Database: pir45  
1 ANNO1  
2 ANNO2  
3 ANNO3  
4 UNANNO1  
5 UNANNO2  
6 UNANNO3  
7 UNANNO4  
8 UNANNO5  
9 UNANNO6  
10 UNANNO7  
11 UNREV1  
12 UNREV2

Statistics: Mean 50.575; Variance 117.403; scale 0.431

Pred. No. is the number of results predicted by chance to  
have a  
score greater than or equal to the score of the result  
being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	%	Query				
No.	Score	Match	Length	DB	ID	Description
Pred. No.						

-----							
1	5765	100.0	771	10	D49423	semaphorin III precu	
0.00e+00	2	5278	91.6	772	9	A49069	collapsin - chicken
0.00e+00	3	917	15.9	730	9	JH0798	fasciclin IV precurs
8.14e-133	4	804	13.9	656	9	B49423	semaphorin I - fruit
7.38e-113	5	799	13.9	724	9	C49423	semaphorin II precur
5.59e-112	6	778	13.5	711	12	A49423	semaphorin I precurs
2.73e-108	7	158	2.7	441	6	S29921	hypothetical protein
1.46e-06	8	158	2.7	403	6	E42521	A39R protein - vacci
1.46e-06	9	147	2.5	295	6	JQ1775	SalL9R protein - vac
3.85e-05	10	116	2.0	122	6	H36852	A43R protein - vario
2.01e-01	11	116	2.0	122	6	JQ1845	14R protein - variol
2.01e-01	12	114	2.0	2733	4	S15760	RNA-directed RNA pol
3.35e-01	13	114	2.0	1174	4	A39927	RNA-directed RNA pol
3.35e-01	14	110	1.9	2731	1	VFIHJH	RNA-directed RNA pol
9.15e-01	15	110	1.9	501	1	PWPMA	H+-transporting ATP
9.15e-01	16	108	1.9	1260	6	S05479	neural cell adhesion
1.50e+00	17	108	1.9	485	4	A44279	spheroidin - Amsacta
1.50e+00	18	108	1.9	648	1	NPVZCP	nucleoside-triphosph
1.50e+00	19	107	1.9	1255	6	S17655	neural cell adhesion
1.91e+00	20	107	1.9	1259	6	S36126	neural cell adhesion
1.91e+00	21	104	1.8	469	9	A25376	SPS2 protein - yeast
3.94e+00	22	103	1.8	254	9	C42691	fibroblast growth fa
4.99e+00	23	103	1.8	713	9	S42803	fibroblast growth fa
4.99e+00	24	102	1.8	823	4	B35963	protein-tyrosine kin
6.32e+00	25	102	1.8	824	12	S24108	protein-tyrosine kin
6.32e+00	26	102	1.8	873	9	B41054	fasciclin II PI-link

6.32e+00	27	102	1.8	813	4	S25060	fibroblast growth fa
6.32e+00	28	102	1.8	806	4	A35963	protein-tyrosine kin
6.32e+00	29	102	1.8	811	9	A41054	fasciclin II, transm
6.32e+00	30	102	1.8	174	6	JQ1611	nonstructural protei
6.32e+00	31	102	1.8	538	6	S22167	neural cell adhesion
6.32e+00	32	102	1.8	174	3	JQ1564	nonstructural protei
6.32e+00	33	102	1.8	813	4	A49123	fibroblast growth fa
6.32e+00	34	101	1.8	705	9	B42691	fibroblast growth fa
7.99e+00	35	101	1.8	682	4	A35969	heparin-binding grow
7.99e+00	36	101	1.8	822	4	A45081	fibroblast growth fa
7.99e+00	37	101	1.8	822	4	A41794	keratinocyte growth
7.99e+00	38	101	1.8	822	1	TVHUF2	fibroblast growth fa
7.99e+00	39	101	1.8	769	4	S16236	fibroblast growth fa
7.99e+00	40	101	1.8	821	4	A42691	fibroblast growth fa
7.99e+00	41	100	1.7	490	4	A32140	steroid 15beta-monoo
1.01e+01	42	100	1.7	1898	10	A45973	trichohyalin - human
1.01e+01	43	99	1.7	811	7	S39901	nwsA protein - Brady
1.27e+01	44	99	1.7	811	7	S32683	nwsA protein - Brady
1.27e+01	45	99	1.7	197	9	A29648	female-specific tran
1.27e+01							

# ALIGNMENTS

```

RESULT      1
ENTRY       D49423      #type complete
TITLE       semaphorin III precursor - human
ORGANISM    #formal name Homo sapiens #common name man
DATE        06-Jan-1995 #sequence_revision 06-Jan-1995
#text_change

          27-Jan-1995
ACCESSIONS  D49423
REFERENCE   A49423
            #authors   Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.
            #journal    Cell (1993) 75:1389-1399

```

```

#title      The Semaphorin genes encode a family of
transmembrane and secreted growth cone guidance molecules.
#accession  D49423
##status    preliminary
##molecule_type mRNA
##residues  1-771 ##label KOL
##cross-references GB:L26081
##note      nucleotide sequence not given
SUMMARY     #length 771 #molecular-weight 88889 #checksum
6249

```

```

DB 10; Score 5765; Match 100.0%; QryMatch 100.0%; Pred.
No. 0.00e+00;
Matches 771; Conservative 0; Mismatches 0; Indels
0; Gaps 0;

```

```

RESULT      2
ENTRY        A49069 #type complete
TITLE        collapsin - chicken
ORGANISM     #formal name Gallus gallus #common name chicken
DATE         07-Apr-1994 #sequence_revision 07-Apr-1994
#text_change

```

```

07-Apr-1994
ACCESSIONS   A49069
REFERENCE    A49069
#authors     Luo, Y.; Raible, D.; Raper, J.A.
#journal     Cell (1993) 75:217-227
#title       Collapsin: a protein in brain that induces the
collapse and

```

```

paralysis of neuronal growth cones.
#accession   A49069
##status     preliminary; not compared with conceptual
translation
##molecule_type mRNA
##residues   1-772 ##label LUO
##cross-references GB:U02528
SUMMARY      #length 772 #molecular-weight 88867 #checksum
9712

```

```

DB 9; Score 5278; Match 88.5%; QryMatch 91.6%; Pred.
No. 0.00e+00;
Matches 683; Conservative 64; Mismatches 24; Indels
1; Gaps 1;

```

```

Db          1
mgwlgiallslgvllagrvcqhvknnvprrlklsykemlesnnivnfnglansssyhtf 60
      |||| | | |||::| | |: |||||
|||||
Qy          1
MGWLTRIVCLFWGVLLTARANYQNGKNNVPRLKLSYKEMLESNNVITFNGLANSSSYHTF 60

```

```

Db          61
lldeersrlyvgakdhifsfnlvnikeyqkivwpvshsrrdeckwagkdilreanfikhv 120

```



```

      |||
:|||||:|||||:
Qy      61
LLDEERSRLYVGAKDHIFSFDLVNIKDFQKIVWPVSYTRRDECKWAGKDILKECANFIKV 120

Db      121
lktynqthlyacgtgafhpmctyievgshpednifirmedshfengrgkspydpklitasl 180
      |||
|||||:|:|||||
Qy      121
LKAYNQTHLYACGTGAFHPICTYIEIGHHPEDNIFKLENSHFENGRGKSPYDPKLLTASL 180

Db      181
lvdgelysgtaadfmgrdfaifrtlghhhpirteqhdsrwlndprfisahlipesdnped 240

|:|||||:|||||:|||||
Qy      181
LIDGELYSGTAADFMGRDFAIFRTLGHHPHPIRTEQHDSRWLNDPKFISAHLISESDNPED 240

Db      241
dkiyfffrenaidgehtgkatharigqickndfgghrslvnkwtflkarlicsvpgpng 300

||:|||||:|||||:|||||
Qy      241
DKVYFFFRENAIDGEHSGKATHARIGQICKNDFGGHRSLVNKWTFLKARLICSVPGPNG 300

Db      301
idthfdelqdvflmnskdpknpivygvfttssnifkgsavcmysmtdvrrvflgpyahrd 360
      |||
|||||:|||||:|||||
Qy      301
IDTHFDELQDVFLMNFKDPKNPVYGVFTTSSNIFKGSAVCMYSMSDVRRVFLGPYAHRD 360

Db      361
gpnysqlwvpyqgrvpyprpgtcpsktfggfdstkdlpdevitfarshpamynpvfpinsrp 420

|||||:|||||:|||||:|
Qy      361
GPNYQWVPYQGRVPYPRPGTCPSKTFGGFDSTKDLRDDVITFARSHPAMYNPVFPMNNRP 420

Db      421
imiktdvdyqftqivvdrvdaedggydvmfigtdigtvlkvvsipketwheleevlleem 480
      |:||||:|||||:|||||
:|||||
Qy      421
IVIKTDVNYQFTQIVVDRVDAEDGQYDVMFIGTDVGTVLKVVSIPKETWYDLEEVILLEEM 480

Db      481
tvfreptvisamkistkqqglyigsatgvsqplhrcdvygkacaeccclardpypcawdgs 540
      |||
|:|||||:|:|||||:|||||
Qy      481
TVFREPTAISAMELSTKQQQLYIGSTAGVAQLPLHRCDIYGKACAECCCLARDPYCAWDGS 540

```



the developing

central nervous system.

KEYWORDS glycoprotein; transmembrane protein

FEATURE

1-22 #domain signal sequence #status predicted  
#label SIG\  
23-730 #product fasciclin IV #status predicted  
#label MAT\  
23-627 #domain extracellular #label EXT\  
628-652 #domain transmembrane #label TMM\  
653-730 #domain intracellular #label INT\  
44,71,163,267,360,  
539 #binding\_site carbohydrate (Asn) (covalent)  
#status

predicted

SUMMARY #length 730 #molecular-weight 81214 #checksum  
5881

DB 9; Score 917; Match 34.8%; QryMatch 15.9%; Pred.  
No. 8.14e-133;  
Matches 174; Conservative 121; Mismatches 167; Indels  
38; Gaps 27;

Db 42  
lgneshkdhfkllekdhnsllvgarnivynislrdltefteqriewhssgahrelcylkg 101  
|:| | | ||: ::: | ||:: :::: | :: :| |:| | |::|:

| |  
Qy 51  
LANSSSYHTF-LLDEERSRLYVGAKDHIFSFDLVNIKDF--QKIVWPVSYTRRDECKWAG 107

Db 102  
kse-ddcqnyirvlakidddrvlicgtnaykplcrhyal--k-dgd-yvvek-eye-grg 154  
| :| |:|:| | :: :: ||| |: |:| : : : :|:

:| |||  
Qy 108  
KDILKECANFIKVLKAYNQTHLYACGTGAFHPICTYIEIGHHPEDNIFKLENSHFENGRG 167

Db 155  
lcpfdpdhnstaiysegqlysatvadfsqtdpliyr--g---plrtersdlkqlnapnfv 209  
|:| | :::: |:|:|:| ||| | | |:| | |:|:|:| | : ||

|:|:  
Qy 168  
KSPYDPKLLTASLLIDGELYSGTAADFMRDFAIFRTLGHHPHPIRTEQHDSRWLNDPKFI 227

Db 210  
nt-m--ey-n---dfiffffretaveyincgkaiysrvarvckhdkggphqfgdrwtsfl 262  
:: : | | | :|:|:| |:: : ||| |:|:|:|:| | | : :

:|:|:|  
Qy 228  
SAHLISESDNPEDDKVYFFFRENAIDGEGHSGKATHARIGQICKNDFGGHRSLVKNWTTFL 287

Db 263  
ksrlncsvpgdypfyfneiqtstsiiegnyggqvekliygvtfttpvnsiggsavcafsmk 322  
|:| | |||| : :: |:| |: : :|:|:|:|:| | : | ||||

:||  
Qy 288  
KARLICSVPGPNGID-THFDELQDVFLMNFKDPKNPVVYGVFTTSSNIFKGSAVCMYSMS 346

Db 323  
silesfdgpfkeqetmnsnwlavpslkvpeprpgqc---v--n-ds-rtlpdvsvnfvks 375  
: | ||: :: | |:: : :|| |||| | || : ||| :

| :|  
Qy 347  
DVRRVFLGPYAHRDGPNYQWVPYQG-RVPYPRPGTCPSKTFGGFDSTKDL PDDVITFARS 405

Db 376  
htmldeavpafftrpilirislqyrftkiavdqqvrtpdgkaydvlfigtddgkvikaln 435  
| | ::| : |||:: : |::| | ||: | : || |||:|||| |

|:| ::  
Qy 406  
HPAMYNPVFPMNNRPIVIKTDVNYQFTQIVVDR-VDAEDGQ-YDVMFIGTDVGTVLKVVS 463

Db 436  
sasfdssdtvdsvvieelqvlppgvpvknlyvvrmdgddsklvvvsddeilaiklhrcgs 495  
: :: :: |::||: |: | : : : : | : | : :

||||  
Qy 464  
IPK-ETWYDLEEVLLEEMTVFR-E-PT-AISAMELSTKQQQLYIGSTAGVAQLPLHRCDI 519

Db 496 dkitncrecvslqdpypcawd 515  
| || :|||||||

Qy 520 YG-KACAECCCLARDPYCAWD 538

Search completed: Wed Nov 1 08:32:08 1995  
Job time : 67 secs.

Release 2.1D John F. Collins, Biocomputing Research Unit.

Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Nov 1 08:30:06 1995; MasPar time 16.72 Seconds

707.202 Million cell

updates/sec

Tabular output not generated.

Title: >US-08-121-713B-54

Description: (1:771) from US08121713B.pep

Perfect Score: 5765

Sequence: 1

MGWLTRIVCLFWGVLLTARA.....ENKKGRNRRTHEFERAPRSV 771

Scoring table: PAM 150

Gap 11

Searched: 43470 seqs, 15335248 residues

Database: swiss-prot31

1 SPT1  
2 SPT2  
3 SPT3  
4 SPT4  
5 SPT5  
6 SPT6  
7 SPT7  
8 SPT8

Statistics: Mean 52.613; Variance 88.879; scale 0.592

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query	%				
No.	Score	Match	Length	DB	ID	Description

Pred. No.

-----  
-----

1	158	2.7	403	7	VA39_VACCC	PROTEIN A39.	
9.12e-10	2	158	2.7	441	7	VA39_VACCV	PROTEIN A39.
9.12e-10	3	114	2.0	2733	6	RRPB_CVMA5	RNA-DIRECTED RNA POLY
1.21e-02	4	110	1.9	501	1	ATPA_PEA	ATP SYNTHASE ALPHA CH
4.50e-02	5	110	1.9	2731	6	RRPB_CVMJH	RNA-DIRECTED RNA POLY
4.50e-02	6	108	1.9	1260	1	CAML_MOUSE	NEURAL CELL ADHESION
8.59e-02	7	108	1.9	648	5	NTP1_CBEPV	NUCLEOSIDE TRIPHOSPHA
8.59e-02	8	107	1.9	1259	1	CAML_RAT	NEURAL CELL ADHESION
1.18e-01	9	104	1.8	469	7	SPS2_YEAST	SPORULATION-SPECIFIC
3.03e-01	10	102	1.8	823	2	CEK3_CHICK	TYROSINE KINASE RECEP
5.61e-01	11	102	1.8	174	8	VNSC_PHODV	NONSTRUCTURAL PROTEIN
5.61e-01	12	102	1.8	811	3	FS22_DROME	FASCICLIN II, PHOSPHA
5.61e-01	13	102	1.8	873	3	FS21_DROME	FASCICLIN II, MEMBRAN
5.61e-01	14	102	1.8	806	2	CEK2_CHICK	TYROSINE KINASE RECEP
5.61e-01	15	102	1.8	813	3	FGR2_XENLA	FIBROBLAST GROWTH FAC
5.61e-01	16	101	1.8	654	1	BFR2_HUMAN	FIBROBLAST GROWTH FAC
7.61e-01	17	101	1.8	821	3	FGR2_HUMAN	FIBROBLAST GROWTH FAC
7.61e-01	18	100	1.7	490	2	CPCC_RAT	CYTOCHROME P450 IIC12
1.03e+00	19	100	1.7	1898	7	TRHY_HUMAN	TRICHOHYALIN.
1.03e+00	20	99	1.7	197	7	TRSF_DROME	FEMALE-SPECIFIC TRANS
1.39e+00	21	98	1.7	1088	6	RRPO_ROTGP	RNA-DIRECTED RNA POLY
1.86e+00	22	98	1.7	707	4	KGFR_MOUSE	KERATINOCYTE GROWTH F
1.86e+00	23	98	1.7	1365	4	KRE5_YEAST	KILLER TOXIN-RESISTAN
1.86e+00	24	97	1.7	701	7	TP20_YEAST	PROTEIN TRANSPORT PRO
2.49e+00	25	96	1.7	859	3	ENV_EIAV3	ENV POLYPROTEIN PRECU
3.33e+00	26	96	1.7	160	5	NCAP_BEV	NUCLEOCAPSID PROTEIN.
3.33e+00	27	96	1.7	239	6	RPSE_BACSU	RNA POLYMERASE SIGMA-
3.33e+00							

28	95	1.6	761	5	NCA2_HUMAN	NEURAL CELL ADHESION
4.44e+00						
29	95	1.6	761	6	RIR1_SALTY	RIBONUCLEOSIDE-DIPHOS
4.44e+00						
30	95	1.6	393	6	RPOC_HALMO	DNA-DIRECTED RNA POLY
4.44e+00						
31	95	1.6	1407	7	TRHY_RABIT	TRICHOHYALIN.
4.44e+00						
32	95	1.6	503	1	ATPA_SYNPI	ATP SYNTHASE ALPHA CH
4.44e+00						
33	95	1.6	859	3	ENV_EIAV1	ENV POLYPROTEIN PRECU
4.44e+00						
34	94	1.6	114	4	HV2A_RABIT	IG HEAVY CHAIN V-A2 R
5.90e+00						
35	94	1.6	1098	6	RPOP_MAIZE	PROBABLE DNA-DIRECTED
5.90e+00						
36	94	1.6	797	5	PHSM_ECOLI	MALTODEXTRIN PHOSPHOR
5.90e+00						
37	94	1.6	1115	5	NCA1_MOUSE	NEURAL CELL ADHESION
5.90e+00						
38	94	1.6	725	5	NCA3_MOUSE	NEURAL CELL ADHESION
5.90e+00						
39	94	1.6	1031	4	KINH_STRPU	KINESIN HEAVY CHAIN.
5.90e+00						
40	94	1.6	182	8	YANC_PSEAE	HYPOTHETICAL 19.8 KD
5.90e+00						
41	94	1.6	513	5	PHSL_DESBA	PERIPLASMIC [NIFESE]
5.90e+00						
42	93	1.6	382	5	MSN1_YEAST	MSN1 PROTEIN (MULTICO
7.81e+00						
43	93	1.6	859	3	ENV_EIAV2	ENV POLYPROTEIN PRECU
7.81e+00						
44	92	1.6	156	8	Y01F_BPT4	HYPOTHETICAL 18.3 KD
1.03e+01						
45	92	1.6	68	4	HSP1_TACAC	SPERM PROTAMINE P1 (C
1.03e+01						

#### ALIGNMENTS

```

RESULT      1
ID   VA39 VACCC      STANDARD;      PRT;      403 AA.
AC   P21062;
DT   01-FEB-1991 (REL. 17, CREATED)
DT   01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT   01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE   PROTEIN A39.
GN   A39R.
OS   VACCINIA VIRUS (STRAIN COPENHAGEN).
OC   VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE;
CHORDOPOXVIRINAE;
OC   ORTHOPOXVIRUSES.
RN   [1]
RP   SEQUENCE FROM N.A.

```

RM 91021027  
 RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW  
 J.P.,  
 RA PAOLETTI E.;  
 RL VIROLOGY 179:247-266(1990).  
 RN [2]  
 RP COMPLETE GENOME.  
 RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW  
 J.P.,  
 RA PAOLETTI E.;  
 RL VIROLOGY 179:517-563(1990).  
 DR EMBL; M35027; PXVACCG.  
 DR PIR; E42521; E42521.  
 SQ SEQUENCE 403 AA; 45741 MW; 907305 CN;

DB 7; Score 158; Match 32.7%; QryMatch 2.7%; Pred. No.  
 9.12e-10;  
 Matches 37; Conservative 29; Mismatches 33; Indels  
 14; Gaps 10;

Db 170  
 dkvyilftd-tigskrivkipy--iaqmclndeggpsslsshrwtflkvelecdi---d 223  
 ||||::| : :| : : | |::| || || || ::|:|||| | |  
 : :  
 Qy 241  
 DKVYFFFRENAIDGESHGKATHARIGQICKNDFGGHRSLV-NKWTTFCLKARLICSVPGPN 299

Db 224 gr-s-yrqi--ihsrtiktdndtily-vffdspys-k-salctysmntikqsf  
 269  
 | : : :: : :| ::::| || |: | ||:| |||: ::| |  
 Qy 300 GIDTHFDELQDVFLMNFKDPKNPVVYGVFTTSSNIFKGSAVCMYSMSDVRVVF  
 352

Search completed: Wed Nov 1 08:30:41 1995  
 Job time : 35 secs.